Orbit complexity, initial data sensitivity and weakly chaotic dynamical systems

Stefano Galatolo

Dipartimento di Matematica

and

Centro Interdisciplinare per lo Studio dei Sistemi Complessi,
Università di Pisa
Via Buonarroti 2/a, Pisa, Italy
e-mail: galatolo@dm.unipi.it

1st February 2008

Abstract

We give a definition of generalized indicators of sensitivity to initial conditions and orbit complexity (a measure of the information that is necessary to describe the orbit of a given point). The well known Ruelle-Pesin and Brin-Katok theorems, combined with Brudno's theorem give a relation between initial data sensitivity and orbit complexity that is generalized in the present work. The generalized relation implies that the set of points where the sensitivity to initial conditions is more than exponential in all directions is a 0 dimensional set. The generalized relation is then applied to the study of an important example of weakly chaotic dynamics: the Manneville map.

1   Introduction

When we think about a chaotic system there are two things that we expect to happen:

- the behavior of the system is unpredictable and complex to be described

- small differences on initial conditions leads to big differences in the evolution of the system.

1

A rigorous measure of initial conditions sensitivity is not difficult to formalize, it leads to the definition of Lyapunov exponents (see e.g. [26]) or to the more general Brin-Katok [9] local entropy.

A measure of the complexity of the behavior of a system is less simple to formalize and in the case of dynamical systems it was given by Brudno [10]. A set of strings is associated by a certain construction to the orbit of a point and then the complexity of the orbit is defined by the Algorithmic Information Content (section 2) of the associated strings. The complexity of an orbit is then a measure of the amount of information that is necessary to describe the orbit.

If dynamics are ergodic and lie on a compact space, it can be proved that the entropy of the system is almost everywhere equal to the orbit complexity. In other words, if such a system has positive entropy, then for a.e. x the algorithmic information that is necessary to describe n steps of the orbit of x increases linearly with n and the proportionality factor is the entropy of the system. This also implies that if a system has an invariant measure μ, its entropy is equal to the mean value of the orbit complexity with respect to μ [4].

It is known that if a system has positive entropy the typical speed of separation of nearby starting orbits is exponential (roughly speaking $\Delta x(t) \simeq \Delta x(0) \, 2^{\lambda t}$). The speed of separation of exponentially divergent orbits is characterized by the number λ. Lyapunov exponents and Brin-Katok [9] local entropy are real valued indicators of the speed of exponential separation of the orbits (the number λ on the exponent of the above formula). The work of Ruelle-Pesin [26] and Brin-Katok shows (under some assumption on the dynamical system) that their indicators are almost everywhere equal to the entropy of the system. In other words: in an ergodic, compact dynamical system the indicator of (exponential) initial data sensitivity is a.e. equal to the entropy which is a.e. equal to the orbit complexity. Thus there is a relation between initial data sensitivity and orbit complexity.

This implies that in the compact case orbit complexity and instability are both faces of the same coin. In the case of compact dynamical systems this motivates the general Ford's claim [15] that orbit complexity was a synonym of chaos. We remark that if the space is not compact there are examples ([4], [17]) of systems with no sensitivity to initial conditions and hight orbit complexity.

In [17] another definition of orbit complexity is given by the use of computable structures, the definition is given by a different approach involving some constructivity concept. This definition is equivalent to Brudno's one if the dynamics lie on a compact space. Replacing compactness with constructivity allows a more powerful investigation of the relation between orbit complexity and chaos in the non compact

2

case.

In many examples of dynamical systems however the entropy could be 0, the speed of separation of nearby starting trajectories could be less than exponential and the increasing of the information contained in n step of the orbit could be less than linear. This is the case of the so called *Weakly Chaotic Dynamics.*

The study of weakly chaotic dynamics was discovered to be important for application purposes. There are connections with many physical and economic phenomena: self organized criticality, the so called chaos threshold, the sporadic dynamics, the anomalous diffusion processes and many others (see for example [5],[12],[32],[19],[24]). In these examples of weakly chaotic dynamics the traditional indicators of chaos (K.S. entropy, Lyapunov exponents, Brudno's orbit complexity) vanishes. These indicators are not able to distinguish between all the various cases of weakly chaotic dynamics. They see all them as trivial dynamical systems.

Some definitions of generalized entropy have been already proposed in literature (see e.g. [30],[24],[32] for definitions about dynamical systems). One of the most fruitful was the one given by Tsallis that found a great variety of applications: long range interacting systems, self gravitating systems, quantum mechanics, social phenomena and many others (see
http://tsallis.cat.cbpf.br/biblio.htm for an updated bibliography of related topics). In the literature relations between Tsallis entropy and initial data sensitivity have been proved ([32],[21]), indicating that the main field of application of Tsallis entropy in dynamical systems is the case of power law initial data sensitivity (if two points starts at distance $\Delta x(0)$ then $\Delta x(t) \sim \Delta x(0)t^{\alpha}$).

The aim of this work is to give a very general definition of indicators of initial data sensitivity and orbit complexity and to prove a relation between them that generalizes the above stated "orbit complexity=sensitivity to initial conditions". This will have some interesting corollary as a consequence (Theorem 42) and finally we will see some applications of our formulas to some example of weakly chaotic dynamics.

The indicators we will define will have values in a totally ordered set $\mathcal{R}$ which is constructed in section 3 by the use of the non standard analysis. In section 2 we add an elementary introduction to non standard analysis, so that the paper is self contained. The ordered space $\mathcal{R}$ will contain (in some way) a representative of all the asymptotic behaviors for $n \to \infty$ of sequences of reals $a_n : \mathbf{N} \to \mathbf{R}$ (for example the various orders of infinity and infinitesimal will correspond to some element of $\mathcal{R}$). In this way we include in one definition all the possible

asymptotic behaviors of initial data sensitivity, orbit complexity or entropy, and so on.

We clarify this by an example. Let us consider the classical definition of topological entropy for a compact dynamical system (X, T). We recall and comment the original definition. If $x, y \in X$ let us say that x, y are (n, ϵ) separated if $d(T^k(x), T^k(y)) > \epsilon$ for some $k \in \{0, ..., n-1\}$. If $d(T^k(x), T^k(y)) \leq \epsilon$ for each $k \in \{0, ..., n-1\}$ then x, y are said to be (n, ϵ) near. That is: two points are separated if they give rise to substantially different orbits. A set $E \subset X$ is called (n, ϵ) separated if $\forall x, y \in E, x \neq y$ then x, y are (n, ϵ) separated. Let us consider

$$s(n, \epsilon) = max\{card(E) : E \subset X \ is \ (n, \epsilon) - separated\}$$

The number $s(n, \epsilon)$ measures the number of substantially different n steps orbits that appear in (X, T).

A chaotic D.S. will have more and more possible different orbits as n increases. The idea of the definition is that in the more chaotic D.S. (where the entropy will be higher) the cardinality $s(n, \epsilon)$ increases more quickly as n increases.

The remaining part of the definition is a way to output a real number from this idea. The number will be a measure of the speed of exponential increasing of the cardinality $s(n, \epsilon)$ as n increases. We define

$$h(T, \epsilon) = limsup_{n \to \infty} \frac{\log s(n, \epsilon)}{n}$$

and then the topological entropy of T is defined as

$$h_{top}(T) = lim_{\epsilon \to 0} h_K(T, \epsilon).$$

The logarithm in the definition is taken because $s(n, \epsilon)$ is expected to increase exponentially $(s(n, \epsilon) \simeq 2^{h(T,\epsilon)n})$. This is a very important case; the case of strongly chaotic dynamic. However, it is worth remarking that there are examples of chaotic topological dynamical systems with 0 entropy (see [29] for example).

When $s(n, \epsilon)$ increases less than exponentially the value of $h(T, \epsilon)$ will be zero. If we want to state a definition of generalized topological entropy that is sensitive to the cases when $s(n, \epsilon)$ increases as a power law we could define ("a la Tsallis" to some extent)

$$h_q(T, \epsilon) = limsup_{n \to \infty} \frac{S_q(s(n, \epsilon))}{n}$$

where $S_q : R \to R$ is

$$S_q(W) = \frac{W^{\frac{1}{1-q}} - 1}{1 - q}.$$

4

In this definition the parameter q plays a role similar to the Hausdorff dimension. Each dynamical system will have a special value of q such that $S_q(s(n, \epsilon))$ increases linearly and allows a nontrivial value of h_q. The special q of each D.S. will be an indicator of the type of chaotic behavior of the system under consideration.This definition will classify the various cases of power law increasing of $s(n, \epsilon)$ and the exponential one (when $q = 1$) but it makes no difference between exponential and stretched exponential[1] (both $q = 1$) and between constant and logarithm ($q = 0$).

We shall see later with the help of examples that the asymptotic behavior of the measures under study exhibit a large variety of different cases. So we need the more possible general definition. This will be done with a definition with values in $\mathcal{R}$. Moreover in this general setting we can prove very general theorems that will have nontrivial meaning in all this variety of examples. We want to remark that the use of the language of the non standard analysis comes about naturally when we want to consider the asymptotic behavior of a sequence of reals. On the other hand, Benci [7] gives an elementary but rigorous approach to the nonstandard analysis which we summarize in a page in section 2.3. This approach is very simple and does not require any deep tool of logic. So we think that our definitions can be easily understood by readers with no experience in this subject.

In section 3 we define the notions of generalized initial data sensitivity at a point x and the notion of generalized complexity of the orbit of x. For the definition of constructivity and orbit complexity we introduce the notion of computable structure on a metric space (Section 2.2). Then, under the assumption that the system is constructive we prove a relation (Theorem 40) between sensitivity and orbit complexity that is a quantitative and rigorous version of the following statement: the asymptotic behavior of the quantity of information that is necessary to reconstruct the orbit of x depends on the initial data sensitivity at x and on the 'complexity' of the point x in X. We remark that theorem 40 is general, it holds even in infinite dimensional spaces, provided that the maps we consider are constructive.

This relation will have as a consequence that the points with upper initial data sensitivity ($R(x)$ in Section 3.1) strictly greater than the exponential one are a 0 dimensional set (Theorem 42).

We remark that our relations are pointwise. We make no use of invariant measures that in many cases of weakly chaotic dynamics are trivial or very complicated (with multifractal support). Instead sometime we make use of the natural measure that can be defined on the metric space X: the Hausdorff measure.

---

[1] $s(n, \epsilon) \simeq 2^{cn^\alpha}$.

In section 4 we apply the the main results of section 3 to give a rigorous estimation of the orbit complexity in an important example of weakly chaotic dynamic: the Manneville maps.

2 AIC, computable structures and non standard analysis

2.1 Algorithmic Information Theory

In this section we give an introduction to algorithmic information theory. The introduction will be informal, to help the reader that is not familiar with recursion theory to understand the paper. A more detailed exposition of algorithmic information theory can be found in [34] or [11].

Let us consider the set $\Sigma = \{0, 1\}^*$ of finite (possibly empty) binary strings. If s is a string we define $|s|$ as the length of s.

Let us consider a Turing machine (a computer) C: by writing $C(p) = s$ we mean that C starting with input p (the program) stops with output s (C defines a partial recursive function $C : \Sigma \to \Sigma$). If the input gives a never ending computation the output (the value of the recursive function) is not defined. If $C : \Sigma \to \Sigma$ is recursive and its value is defined for all the input strings in Σ (the computation stops for each input) then we say that C is a *total* recursive function from Σ to Σ. The algorithmic information content of a string will be the length of the shortest program that outputs the string.

Definition 1 *The Kolmogorov complexity or algorithmic information content of a string s given C is the length of the smallest program p giving s as the output:*

$$K_C(s) = \min_{C(p)=s} |p|$$

if s is not a possible output for the computer C then $K_C(s) = \infty$.

For example by this definition we see that the algorithmic information content (A.I.C.) of a $2n$ bits long periodic string

$$s ='' 1010101010101010101010...''$$

is small because the string is output of a shortest program:
$$repeat \; n \; times \; (write \; (``10''))$$
the AIC of the string s is then less or equal than $log(n) + Constant$ where $log(n)$ bits are sufficient to code "n" and the constant represents

the length of the code for the computer C representing the instructions "repeat...". As it is intuitive the information content of a periodic string is very poor. On the other hand each n bits long string

$$s' =^{''} 1010110101010010110...''$$

is output of the trivial program

$$write(^{''}1010110101010010110...^{''})$$

this is of length $n + constant$ this implies that the AIC of each string is (modulo a constant which depends on the chosen computer C) less or equal than its length.

Until this point the algorithmic information content of a string depends on the choice of the computer C. We will see that there is a class of computers that allows an "almost" universal definition of algorithmic information content of a string: if we consider computers from this class the A.I.C. of a string will be defined independently of the computer up to a constant. In order to define such a class of universal computers we give some notations that are necessary to work with strings: there is a correspondence $c : \Sigma \to \mathbf{N}$ from the set Σ and the set $\mathbf{N}$ of natural numbers

$$\emptyset \to 0, \quad 0 \to 1, \quad 1 \to 2, \quad 00 \to 3, \quad 01 \to 4, \quad 10 \to 5, ...$$

This correspondence allows us to interpret natural numbers as strings and vice-versa when it is needed. We remark that $|s| \leq log(c(s)) + 1$. [2]

If s is a string with $|s| = n$ we denote by $s\,\hat{}\,$ the string $s_0 s_0 s_1 s_1 ... s_{n-1} s_{n-1} 01$. If $a = a_1 ... a_n$ and $b = b_1 ... b_m$ are strings then ab is defined as the string $a_1 ... a_n b_1 ... b_m$. If a and b are strings then $a\,\hat{}\,b$ is an encoding of the couple (a, b). There is an algorithm that getting the string $a\,\hat{}\,b$ is able to recover both the strings a and b. An universal Turing machine intuitively is a machine that can emulate any other Turing machine if an appropriate input is given. We recall that there is a recursive enumeration $A_1, A_2...$ of all the Turing machines.

Definition 2 *A Turing machine $\mathcal{U}$ is said to be universal if for all $m \in \mathbf{N}$ and $p \in \Sigma$ then $\mathcal{U}(c^{-1}(m)\,\hat{}\,p) = A_m(p)$.*

In the last definition the machine $\mathcal{U}$ is universal because $\mathcal{U}$ is able to emulate each other machine A_m when in its input we specify the number m identificating A_m and the program to be runned by A_m. It can be proved that an universal Turing machine exists.

---

[2] In this paper all the logarithms are in base two.

Definition 3 *A Turing machine F is said to be* asymptotically optimal *if for each Turing machine C and each binary string s we have $K_F(s) \leq K_C(s) + c$ where the constant c depends on C and not on s.*

The following proposition can be proved from the definitions

Proposition 4 *If $\mathcal{U}$ is an universal Turing machine then $\mathcal{U}$ is asimptotically optimal.*

This tells us that choosing an universal Turing machine the complexity of a string is defined independently of the given Turing machine up to a constant. For the remaining part of the paper we will suppose that an universal Turing machine $\mathcal{U}(p)$ is chosen once forever.

2.2 Computable Structures, Constructivity

A computable structure on a separable metric space (X, d) is a class of dense immersions $(I : \Sigma \to X)$ of the space of finite strings Σ in the metric space. The immersions are such that the distance d restricted to the points that are images of strings $(x = I(s) : x \in X, s \in \Sigma)$ is a "computable" function. Many concrete metric spaces used in analysis or in geometry have a natural choice of a computable structure. The use of computable structures allows to consider algorithms acting over metric spaces and to define constructive functions between metric spaces, that is, functions such that we can work with by using a finite amount of information. In the following we often will assume that the dynamical systems under our consideration are constructive. All the dynamical system that we can construct explicitely are construcive. From the philosophical point of wiew we think that the assumption of constructivity is not unnatural because even if the maps coming from physical reality were not constructive, the models used to describe such a reality should be constructive (to allow calculations). On the other hand, to add constructivity allows to prove stronger theorems, avoiding pathologies coming from random maps.

An interpretation function is a way to interpret a string as a point of the metric space.

Definition 5 *An interpretation function on (X, d) is a function $I : \Sigma \to X$ such that $I(\Sigma)$ is dense in X.*

A point $x \in X$ is said to be *ideal* if it is the image of some string $x = I(s), s \in \Sigma$. An interpretation is said to be computable if the distance between ideal points is computable with arbitrary precision:

Definition 6 *A computable interpretation function on (X, d) is a function $I : \Sigma \to X$ such that $I(\Sigma)$ is dense in X and there exists a total recursive function $D : \Sigma \times \Sigma \times \mathbf{N} \to \mathbf{Q}$ such that $\forall s_1, s_2 \in \Sigma, n \in \mathbf{N}$:*

$$|d(I(s_1), I(s_2)) - D(s_1, s_2, n)| \leq \frac{1}{2^n}.$$

Two interpretations are said to be equivalent if the distance from an ideal point from the first and a point from the second is computable up to arbitrary precision. For example, the finite binary strings $s \in \Sigma$ can be interpreted as rational numbers by interpreting the string as the binary expansion of a number. Another interpretation can be given by interpreting a string as an encoding of a couple of integers whose ratio gives the rational number. If the last encoding is recursive, the two interpretation are equivalent.

Definition 7 *Let I_1 and I_2 be two computable interpretations in (X, d); we say that I_1 and I_2 are equivalent if there exists a total recursive function $D^* : \Sigma \times \Sigma \times \mathbf{N} \to \mathbf{Q}$, such that $\forall s_1, s_2 \in \Sigma, n \in \mathbf{N}$:*

$$|d(I_1(s_1), I_2(s_2)) - D^*(s_1, s_2, n)| \leq \frac{1}{2^n}.$$

Proposition 8 *The relation defined by definition 7 is an equivalence relation.*

For the proof of this proposition see [17].

Definition 9 *A computable structure $\mathcal{I}$ on X is an equivalence class of computable interpretations in X.*

For example if $X = \mathbf{R}$ we can consider the interpretation $I : \Sigma \to \mathbf{R}$ defined in the following way: if $s = s_1...s_n \in \Sigma$ then

$$I(s) = \sum_{1 \leq i \leq n} s_i 2^{[n/2]-i}. \tag{1}$$

This is an interpretation of a string as a binary expansion of a number. I is a computable interpretation, the computable structure on $\mathbf{R}$ containing I will be called *standard* computable structure. If $r = r_1 r_2...$ is an infinite string such that $lim\frac{K_C(r_1...r_n)}{n} = 1$[3] then the interpretation I_r defined as $I_r(s) = I(s) + \sum r(i)2^{-i}$ is computable but not equivalent to I. I and I_r belongs to different computable structures.

---

[3] such a string exist, see for example [16] theorem 13.

In a similar way it is easy to construct computable structures in $\mathbf{R}^n$ or in separable function spaces codifying a dense subset (for example the set of step functions) with finite strings. We remark as a property of the computable structures that if $B_r(I(s))$ is an open ball with center in an ideal point $I(s)$ and rational radius r and $I(t)$ is another point then there is an algorithm that verifies if $I(t) \in B_r(I(s))$. If $I(t) \in B_r(I(s))$ then the algorithm outputs "yes", if $I(t) \notin B_r(I(s))$ the algorithm outputs "no" or does not stop. The algorithm calculates $D(s,t,n)$ for each n until it finds that $D(s,t,n) + 2^{-n} < r$ or $D(s,t,n) - 2^{-n} > r$, in the first case it outputs "yes" and in the second it outputs "no", if $d(I(s), I(t)) \neq r$ the algorithm will stop and output an answer.

We give a definition of *morphism* of metric spaces with computable structures, a morphism is heuristically a computable function between computable metric spaces.

Definition 10 *If $(X, d, \mathcal{I})$ and $(Y, d', \mathcal{J})$ are spaces with computable structures; a function $\Psi : X \to Y$ is said to be a morphism of computable structures if Ψ is uniformly continuous and for each pair $I \in \mathcal{I}, J \in \mathcal{J}$ there exists a total recursive function $D^* : \Sigma \times \Sigma \times \mathbf{N} \to \mathbf{Q}$, such that $\forall s_1, s_2 \in \Sigma, n \in \mathbf{N}$:*

$$|d'(\Psi(I(s_1)), J(s_2)) - D^*(s_1, s_2, n)| \leq \frac{1}{2^n}.$$

We remark that Ψ is not required to have dense image and then $\Psi(I(*))$ is not necessarily an interpretation function equivalent to J.

Remark 11 *As an example of the properties of the morphisms, we remark that if a map $\Psi : X \to Y$ is a morphism then given a point $x \in I(\Sigma) \subset X$ it is possible to find by an algorithm a point $y \in J(\Sigma) \subset Y$ as near as we want to $\Psi(x)$.*

The procedure is simple: if $x = I(s)$ and we want to find a point $y = J(z_0)$ such that $d'(\Psi(I(s)), y) \leq 2^{-m}$ then we calculate $D^*(s, z, m+2)$ for each $z \in \Sigma$ until we find z_0 such that $D^*(s, z_0, m+2) < 2^{-m-1}$. Clearly $y = J(z_0)$ is such that $d'(\Psi(x), y) \leq 2^{-m}$. The existence of such a z_0 is assured by the density of J in Y. In particular the identity is a morphism. Remark 11 applied to the identity will be used in the proof of lemma 13.

A constructive map is a morphism for which the continuity relation between ϵ and δ is given by a recursive function. The following is in some sense a generalization of the definition of Grzegorczyk, Lacombe (see e.g. [28]) of constructive function.

10

Definition 12 *A function $\Psi : (X, d, \mathcal{I}) \to (Y, d', \mathcal{J})$ between spaces with computable structure $(X, \mathcal{I}), (Y, \mathcal{J})$ is said to be constructive if Ψ is a morphism between the computable structures and it is effectively uniformly continuous, i.e. there is a total recursive function $f : \mathbf{N} \to \mathbf{N}$ such that for all $x, y \in X$ $d(x, y) < 2^{-f(n)}$ implies $d'(\Psi(x), \Psi(y)) < 2^{-n}$.*

If $f(n)$ is recursive and satisfies the hypothesis that, $d(x, y) < 2^{-f(n)}$ implies $d(T(x), T(y)) < 2^{-n}$ then $max(f(n), n)$ is recursive and still satisfies the hypothesis. Then we can suppose that if $f(n)$ is a function of effective continuity then $f(n) \geq n$.

If a map between spaces with a computable structure is constructive then there is an algorithm to follow the orbit each ideal point $x = I(s_0)$.

Lemma 13 *If $T : (X, \mathcal{I}) \to (X, \mathcal{I})$ is constructive, $I \in \mathcal{I}$ then there is an algorithm (a total recursive function) $A : \Sigma \times \mathbf{N} \times \mathbf{N} \to \Sigma$ such that $\forall k, m \in \mathbf{N}, s_0 \in \Sigma$ $d(T^k(I(s_0)), I(A(s_0, k, m))) < 2^{-m}$.*

Proof. Since T is effectively uniformly continuous we define the function $g_k(m)$ inductively as $g_1(m) = f(m)+1$, $g_i(m) = f(g_{i-1}(m)+1)$ where f is the function of effective uniform continuity of T (definition 12). If $d(x, y) < 2^{-g_k(m)}$ then $d(T^i(y), T^i(x)) < 2^{-m}$ for $i \in \{1, ..., k\}$. Let us choose $I \in \mathcal{I}$. We recall that the assumption that T is a morphism implies that there is a recursive function $D^*(s_1, s_2, n)$ such that

$$|D^*(s_1, s_2, n) - d(I(s_1), T(I(s_2)))| < 2^{-n}.$$

Let us suppose that $x = I(s_0)$. Now let us describe the algorithm A: using the function D^* and the function f, A calculates $g_k(m)$ and finds a string s_1 such that $d(I(s_1), T(I(s_0))) < 2^{-g_k(m)}$ as described in remark 11. This is the first step of the algorithm. Now $d(T(I(s_1)), T^2(x)) \leq 2^{-(g_{k-1}+1)}$. We can use D^* to find a string s_2 such that $d(I(s_2), T(I(s_1))) < 2^{-(g_{k-1}+1)}$. By this $d(I(s_2), T^2(x)) \leq 2^{-g_{k-1}}$. This implies that $d(T(I(s_2)), T^3(x)) \leq 2^{-(g_{k-2}+1)}$, then we find s_3 such that $d(I(s_3), T(s_2)) \leq 2^{-(g_{k-2}+1)}$ and so on for k steps. At the end we find a string s_k such that $d(I(s_k), T^k(x)) \leq 2^{-m}$. $\square$

2.3 Non standard analysis

We define the extended real line $\mathbf{R}^*$ to be an ordered field satisfying suitable axioms. The existence of such a field is proved in [7]. $\mathbf{R}^*$ will contain the standard real numbers and other elements representing *infinite* and the *infinitesimal* numbers.

We call Hyperreal Line a field satisfying the following axioms:

Axiom 14 *The set of the hyperreal numbers* $\mathbf{R}^*$ *is an ordered field which contains* $\mathbf{R}$ *as a subfield.*

Axiom 15 *There is a surjective ring homomorphism*

$$J : \mathbf{R}^{\mathbf{N}} \to \mathbf{R}^*$$

associating to each real sequence an hyperreal number.

Intuitively the homomorphism J associates to a sequence of reals its asymptotic behavior. For example if $\lim_{n\to\infty} a_n = 0$ then $J(a_n)$ will be an *infinitesimal* number. Moreover if $a = (a_i)$ and $b = (b_i)$ are two sequences we would like that if $a_i \geq b_i$ for all $i \in \mathbf{N}$ then $J(a) \geq J(b)$. For this reason J is required to satisfy the following monotonicity property

Axiom 16 *If there exists* $k \in \mathbf{N}^+$ *such that*

$$\forall n \in \mathbf{N}^+, \phi_{kn} \geq a$$

with $a \in \mathbf{R}$, *then*

$$J(\phi) \geq a \ .$$

From axiom 16 it follows for example that if $a_i = 2^{-i}$ and $b_i = \frac{1}{i}$ then $0 \leq J(a_i) \leq J(b_i)$. Another consequence of axiom 16 is that for each real sequence x_i $liminf(x_i) \leq J(x_i) \leq limsup(x_i)$.

A field satisfying our axioms exists. As the reader could imagine, it can be constructed from the set of real sequences modulo a suitable equivalence relation ([7]).

2.3.1 Extension of functions, infinite and infinitesimal numbers.

Given any function $f : \mathbf{R} \to \mathbf{R}$ we extend it to a function $f^* : \mathbf{R}^* \to \mathbf{R}^*$ as follows: if a_i is a real sequence and $x = J(a_i) \in \mathbf{R}^*$ we define

$$f^*(x) = f^*(J(a_i)) = J(f(a_i)).$$

Proposition 17 *the definition is well posed i.e.* $J(a_i) = J(b_i)$ *implies that* $J(f(a_i)) = J(f(b_i))$.

The proof of proposition 17 can be found in [7].

As we stated before, in $\mathbf{R}^*$ there are some elements representing the infinite and infinitesimal numbers:

Definition 18 *An hyperreal number ξ is called infinite if $\forall k \in \mathbf{N}$ we have $|\xi| > k$. A number ξ is called infinitesimal if $\forall k \in \mathbf{N}, |\xi| < \frac{1}{k}$. A number ξ is called bounded if $\exists k \in \mathbf{N}, |\xi| < k$.*

For example the reader could verify directly from the axioms that if $a_i : \mathbf{N} \to \mathbf{R}$ is the identity: $a_i = i$ then $J(a_i)$ is infinite and $J(\frac{1}{a_i})$ is infinitesimal.

3 sensitivity and orbit complexity

Now we construct the space $\mathcal{R}$ in which our indicators of orbit complexity and initial data sensitivity will have value.

Definition 19 *If $a, b \in \mathbf{R}^*$ we say that a and b have the same order and write $a \simeq b$ if and only if both $\frac{a}{b}$ and $\frac{b}{a}$ are bounded.*

$\simeq$ it is clearly an equivalence relation. In the following by $[\alpha]$ we will indicate the equivalence class of α.

We now define an ordering relation on the quotient space $\frac{\mathbf{R}^*}{\simeq}$. We say that $[a] \leq [b]^4$ if $\forall x \in [a], \forall y \in [b]$ then or $x \simeq y$ or $x < y$. We remark that the order relation on $\mathbf{R}^*$ is compatible with the equivalence relation $\simeq$, if $a \leq b$ in $\mathbf{R}^*$ then $[a] \leq [b]$ in $\frac{\mathbf{R}^*}{\simeq}$. And thus $\frac{\mathbf{R}^*}{\simeq}$ is totally ordered. The relation $<$ is then defined in the obviuos way as $[a] < [b] \iff [a] \leq [b], [a] \neq [b]$.

$\frac{\mathbf{R}^*}{\simeq}$ contains a representative of all the infinite (infinitesimal) asymptotic behaviors of real sequences. $\frac{\mathbf{R}^*}{\simeq}$ is sometime called the group of orders. The natural projection from $\mathbf{R}^*$ to $\frac{\mathbf{R}^*}{\simeq} : a \to [a]$ allows to forget all the lower order terms in the hyperreal number a: for example if $a_i = i$ as above $J(a_i{}^2 + a_i)$ and $J(a_i{}^2 + \sqrt{a_i})$ belongs to the same class as $J(a_i{}^2)$, in other words $[J(a_i{}^2 + a_i)] = [J(a_i{}^2 + \sqrt{a_i})] = [J(a_i{}^2)]$.

Unfortunately $\frac{\mathbf{R}^*}{\simeq}$ is not complete (as $\mathbf{R}^*$ is not complete). The supremum or the infimum of a sequence in $\frac{\mathbf{R}^*}{\simeq}$ may not exist in $\frac{\mathbf{R}^*}{\simeq}$. For this reason we will consider a space $\mathcal{R}$ which is a completion of $\frac{\mathbf{R}^*}{\simeq}$, the *sup* and *inf* of each sequence in $\frac{\mathbf{R}^*}{\simeq}$ is in $\mathcal{R}$. We now outline a possible construction of a completition of $\frac{\mathbf{R}^*}{\simeq}$, there are other possible costructions. Another possible completition of $\frac{\mathbf{R}^*}{\simeq}$ can be constructed for example by Dedekind sections. We construct $\mathcal{R}$ by quotienting the set of monotone sequences in $\frac{\mathbf{R}^*}{\simeq}$ by a suitable equivalence relation. This will add the supremum to each countable sequence.

---

[4]By an abuse of notations we use the symbol $\leq$ for this ordering relation, this will cause no ambiguity with the ordering relation defined on $\mathbf{R}^*$.

Proposition 20 *There is an ordered space $\mathcal{R}$ such that $\frac{\mathbf{R}^*}{\simeq} \subset \mathcal{R}$ in a natural and order preserving way and if $a_i \in \frac{\mathbf{R}^*}{\simeq}$ is a monotone sequence then $inf(a_i)$ and $sup(a_i)$ are in $\mathcal{R}$.*

Proof. Let us consider the set of monotone sequences in $\frac{(\mathbf{R}^*)}{\simeq}$:

$$\mathcal{A} = \{(a_i) : \mathbf{N} \to \frac{\mathbf{R}^*}{\simeq} \ s.t. \ a_i \ is \ monotone\}$$

in $\mathcal{A}$ the set $\frac{\mathbf{R}^*}{\simeq}$ will be identified with the subset of constant sequences.

We define the ordering relation on $\mathcal{A}$ in the following way $(a_i) < (b_j) \iff \exists M$ such that $\forall n, m$ with $n > M, m > M$ then $a_n < b_m$.

We define the relation $\approx$ in the following way: $(a_i) \approx (b_j)$ if neither $(a_i) < (b_j)$ nor $(b_j) < (a_i)$. $\approx$ is an equivalence relation: $\approx$ is trivially symmetric and reflexive. The transitivity follows from the remark that if a_i is not $< b_j$ then $\forall M \ \exists n, m \ s.t. \ n > M, m > M, a_n \geq b_m$. As it is easy to verify this is a transitive relation (because the sequences in $\mathcal{A}$ are monotone).

The set of equivalence classes $\mathcal{R} = \frac{\mathcal{A}}{\approx}$ is then totally ordered in the same way as before and contains $\frac{\mathbf{R}^*}{\simeq}$ as a subset. Moreover each monotone sequence in $\frac{\mathbf{R}^*}{\simeq}$ has its *sup* and *inf* in $\mathcal{R}$: let us indicate with p the *natural projection* associating to each element in $\mathcal{A}$ its equivalence class. If for example $a = (a_i)$ is a nondecreasing sequence then $inf(a)$ is the equivalence class of the constant sequence $b = (b_j)$ such that $\forall j \ b_j = a_0$ and $sup(a) = p(a)$ (p is the natural projection map as defined above). It is easy to verify that $p(a) > a_j$ for each $j \in \mathbf{N}$ and for each $x \in \mathcal{R} \ s.t. \ \forall j \in \mathbf{N}, x > a_j$ we have $x > p(a)$ or $x \approx p(a)$. $\square$

The set $\mathcal{R}$ may look in some way mysterious and not practical to be used. The reader will see in the examples that the elements of $\mathcal{R}$ we will have as value of our invariants will be classes that can be expressed in an explicit way. For example a possible value of $r(x)$ (definition 24) could be $[J(n^{-\frac{1}{2}})]$ (the class in $\frac{R^*}{\simeq}$ containing the asymptotic behavior of the sequence $a_n = n^{-\frac{1}{2}}$). Since $\frac{R^*}{\simeq}$ is immersed in $\mathcal{R}$ in a natural way we can consider an element of $\frac{R^*}{\simeq}$ as an element of $\mathcal{R}$ without ambiguity. Another possible value of $r(x)$ could be $[J(n^{-\frac{1}{3}})]$ and it is clear that $[J(n^{-\frac{1}{2}})] \leq [J(n^{-\frac{1}{3}})]$ so we can easily compare the values.

The notion of infinite and infinitesimal numbers can be extended to the elements of $\mathcal{R}$: if $v \in \mathbf{R}, v > 0$ then $c = [J(v)]$ is the element of $\mathcal{R}$ corresponding to the class of bounded and not infinitesimal numbers. Moreover: an element $\epsilon \in \mathcal{R}$ is said to be infinitesimal if $\epsilon < c$ and an element $\gamma \in \mathcal{R}$ is said to be infinite if $\epsilon > c$.

Finally we remark that $\mathcal{R}$ is closed by countable inf and sup and the projection p (defined in the proof of the above proposition) can

be extended to a function from the set of monotone sequences in $\mathcal{R}$ to $\mathcal{R}$, associating to a sequence its supremum or infimum according that the sequence is increasing or decreasing.

Proposition 21 *If (a_i) is a monotone sequence in $\mathcal{R}$, then*

$$sup(a_i), inf(a_i) \in \mathcal{R}$$

.

Proof. Let us suppose that (a_i) is a non decreasing sequence, the case where (a_i) is non increasing is analogous. We show that $sup(a_i) \in \mathcal{R}$. If (a_i) is eventually constant the proposition is obvious. If (a_i) is not eventually constant let us consider a subsequence a_{i_k} that is strictly increasing ($\forall k \ a_{i_k} > a_{i_{k-1}}$). The elements $a_{i_k}, a_{i_{k-1}} \in \mathcal{R}$ are equivalence classes of monotone sequences with values in $\frac{R^*}{\simeq}$. Let use consider two of this sequences $(\alpha_n^{k-1})_{n \in \mathbf{N}} : \mathbf{N} \to \frac{R^*}{\simeq}$ such that $p((\alpha_n^{k-1})_{n \in \mathbf{N}}) = (a_{i_{k-1}})$ and $(\alpha_n^k)_{n \in \mathbf{N}} : \mathbf{N} \to \frac{R^*}{\simeq}$ such that $p((\alpha_n^k)_{n \in \mathbf{N}}) = (a_{i_{k-1}})$. Since $(\alpha_n^{k-1}) < (\alpha_n^k)$ then $\exists M_k \in \mathbf{N}$ s.t. $\forall n \ \alpha_n^{k-1} < \alpha_{M_k}^k$. This is true for each choice of the sequences α^k, α^{k-1} in the classes a_k, a_{k-1}, hence what follows does not depends on the choice of α^k and α^{k-1}. Let us now consider the sequence $(\alpha_{M_k}^k)_{k \in \mathbf{N}} : \mathbf{N} \to \frac{R^*}{\simeq}$ this is an increasing sequence and $\forall j < k, \ a_{i_j} < \alpha_{M_k}^k$, then $\forall i \in \mathbf{N}, p(\alpha_{M_k}^k) > a_i$. Now it is easy to see that if b is such that $\forall i_k b > a_{i_k}$ then $b \approx p(\alpha_{M_k}^k)$ or $b > p(\alpha_{M_k}^k)$, then $p(\alpha_{M_k}^k)$ is $sup(a_i)$. $\square$

By the above result we also see that the function $p : \mathcal{A} \to \mathcal{R}$ can be extended to a function $\overline{p} : \{(a_i) : \mathbf{N} \to \mathcal{R} \ s.t. \ a_i \ is \ monotone\} \to \mathcal{R}$ by associating to each sequence its *sup* of *inf* according that the sequence is increasing or decreasing.

3.1   Initial data sensitivity

Let X be a separable metric space and T a function $X \to X$. [5] Let us consider the following set:

$$B(n, x, \epsilon) = \{y \in X : d(T^i(y), T^i(x)) \le \epsilon \ \forall i \ s.t. \ 0 \le i \le n\}.$$

$B(n, x, \epsilon)$ is the set of points "following" the orbit of x for n steps at a distance less than ϵ. When the orbits of (X, T) diverges the set $B(n, x, \epsilon)$ will be smaller and smaller as n increases. The speed of

---

[5]We remark that in this and in the following subsection we do not require that T is continuous.

decreasing of the size of this set considered as a function of n will be a measure of the sensitivity of the system to changes on initial conditions.

Brin and Katok used the set $B(n, x, \epsilon)$ for their definition of local entropy [9]. In their paper the measure of the size of $B(n, x, \epsilon)$ was the invariant measure of the set.

If we are interested to approximate the orbit of x for n steps we are interested to know how close we must approach the initial condition x to ensure that the resulting approximate orbit is close to the orbit of x; another possible measure of the size of $B(n, x, \epsilon)$ is then the radius of the biggest ball with center x contained in $B(n, x, \epsilon)$.

$$r(x, n, \epsilon) = \sup_{B_r(x) \subset B(n,x,\epsilon)} r \quad .$$

Or the radius of the smaller ball that contains $B(n, x, \epsilon)$

$$R(x, n, \epsilon) = \inf_{B_R(x) \supset B(n,x,\epsilon)} R \quad .$$

As said before the generalized initial data sensitivity will be a function associating to a point of X a class in $\mathcal{R}$ indicating how faster orbits coming from a neighborhood of x will diverge. For this purpose we measure how faster $r(x, n, \epsilon)$ decreases as n increases, i.e. we consider the asymptotic behavior of the sequence $r(x, n, \epsilon)$ as n increases. First we define

Definition 22 *We define $r_\epsilon : X \to \frac{R^*}{\simeq}$ as*

$$r_\epsilon(x) = [J(r(x, n, \epsilon))]$$

and
$R_\epsilon : X \to \frac{R^*}{\simeq}$ *as*

$$R_\epsilon(x) = [J(R(x, n, \epsilon))]$$

The following lemma implies that $r_\epsilon(x)$ and $R_\epsilon(x)$ are monotone functions with respect to ϵ.

Lemma 23 *If $\epsilon > \theta$ then $r_\epsilon(x) \geq r_\theta(x)$, $R_\epsilon(x) \geq R_\theta(x)$.*

Proof. Obvious □
For the previous lemma we define the indicator of initial data sensitivity at x by letting ϵ go to 0 as the infimum of the $r_\epsilon(x)$ for $\epsilon \in \mathbf{R}, \epsilon > 0$. This infimum will be in $\mathcal{R}$.

Definition 24 *We define the indicator of initial data sensitivity at x as*

$$r(x) : X \to \mathcal{R}$$

$$r(x) = \inf_{\epsilon \in \mathbf{R}^+} r_\epsilon(x)$$

In the same way we define

$$R(x) : X \to \mathcal{R}$$

$$R(x) = \inf_{\epsilon \in \mathbf{R}^+} R_\epsilon(x).$$

The classical definition of dynamical system sensitive to initial conditions is related to our last definition. To say that a system is sensitive to initial conditions is equivalent to say that there is a δ such that $r_\delta(x)$ is infinitesimal for all the $x \in X$:

Definition 25 *A dynamical system (X, T) is said to have sensitive dependence on initial conditions if there is a δ such that for each $x \in X$ and every neighborhood U of x there is $y \in U$ and $k \in \mathbf{N}$ such that $d(T^k(x), T^k(y)) > \delta$.*

Proposition 26 *A system has sensitive dependence on initial conditions if and only if there is a δ such that $\forall x \in X$ $r_\delta(x)$ is infinitesimal.*

The proof follows directly from the definition of $r_\delta(x)$.

We give some example of different behaviors of $r(x)$ and $R(x)$ in dynamical system over the interval $[0,1]$. *The identity map $T(x) = x$.* In this map $\forall n$ $B(n, x, \epsilon) = \{y \in [0,1], s.t. |y-x| < \epsilon\}$ then if we choose for example $x = \frac{1}{2}$ we have $R(\frac{1}{2}, n, \epsilon) = r(\frac{1}{2}, n, \epsilon) = \epsilon$ i.e. the constant sequence with value ϵ. Then $R_\epsilon(\frac{1}{2}) = r_\epsilon(\frac{1}{2}) = [J(\epsilon)]$ where $J(\epsilon)$ is the number in R^* corresponding to the constant sequence with value ϵ and $R(\frac{1}{2}) = r(\frac{1}{2}) = [J(\epsilon)] = \{x \in R^* \ s.t. \ x \ is \ bounded\}$ i.e. the class containing the numbers in R^* corresponding to the constant sequences. The same arguments can be applied to the irrational translation on $[0,1]$: $T(x) = x + t \ (mod \ 1)$ where $t \notin \mathbf{Q}$ obtaining the same kind of initial data sensitivity as the identity (in effect both the maps are not sensitive to initial conditions).
The one dimensional baker's map $T : [0,1] \to [0,1], T(x) = 2x \ (mod \ 1)$. If we choose for example $x = 0$, we have $B(n, 0, \epsilon) = \{y \in [0,1], 0 \le y \le 2^{-n}\epsilon\}$, $R_\epsilon(0) = r_\epsilon(0) = [J(\epsilon 2^{-n})]$ and $R(0) = r(0) = [J(2^{-n})]$ i.e. the class containing all the bounded multiples of the exponential infinitesimal number.
The piecewise linear map $T : [0,1] \to [0,1]$

$$T(x) = \begin{cases} \frac{\xi_{k-2}-\xi_{k-1}}{\xi_{k-1}-\xi_k}(x-\xi_k)+\xi_{k-1} & \xi_k \leq x < \xi_{k-1} \\ \frac{x-a}{1-a} & a \leq x \leq 1 \end{cases} \quad \text{with } \xi_k =$$

$\frac{a}{(k+1)^{\frac{1}{z-1}}}$, $k \in N, z \in \mathbf{R}$, $z \geq 2$. This is a P.L. version of the Manneville map $T(x) = x + x^z \pmod 1$ (see fig 1), this example will be discussed more deeply in Section 4. In this example any neighborhood of the origin $B_\epsilon = [0, \epsilon)$ is subdivided in a sequence of intervals $A_k = (\xi_k, \xi_{k-1}]$ and if $k > 1$ then $T(A_k) = A_{k-1}$. Let us choose $x = 0$, then $B(n, 0, \xi_k) = [0, \xi_{k+n})$. By this we find $r(0) = [J\left(\frac{a}{(n+k+1)^{\frac{1}{z-1}}}\right)]$ that is a class of infinitesimals corresponding to power law decreasing sequences whit exponent $\frac{-1}{z-1}$. In other words the map $T(x)$ has power law sensitivity to initial condition at the origin. In Section 4 we will see that while the sensitivity to initial condition at the origin is a power law, for almost all other points in $[0, 1]$ we have a stretched exponential sensitivity. This example is important in the applications and will be studied more deeply in section 4.

3.2 orbit complexity

Now we define our indicator of orbit complexity. In the philosophy of the algorithmic information content we define the complexity of the orbit of x as the asymptotic behavior of the quantity of information that is necessary to reconstruct the orbit, i.e. the asymptotic behavior (with respect to the variable n) of the length of the smallest program that can approximate n steps of the orbit with its output (at accuracy ϵ). As before we consider the behavior when n goes to ∞ and the accuracy parameter goes to 0 .

To interpret the output of a calculation which is a finite string as a finite sequence in X let us consider an interpretation function I and a total recursive surjective function.

$$\mathcal{Q} : \Sigma \rightarrow \Sigma^*$$

where Σ^* is the set of finite sequences in Σ. Now let us consider an universal Turing machine $\mathcal{U}$, for each program p we define $U(p) \in X^*$ (the set of finite sequences in X) as

$$U(p) = I(\mathcal{Q}(\mathcal{U}(p)))$$

where I is extended obviously to a map from the space Σ^* to X^*. $U_i(p) \in X$ is defined as the $i-$th point of $U(p)$. With this definition we can interpret the output of a calculation as a finite sequence in X. We remark that given $\mathcal{Q}$ and a sequence of strings $s_1, ..., s_n$ it is possible by an algorithm to find a single string s such that $\mathcal{Q}(s) = (s_1, ..., s_n)$.

Definition 27 *We define the algorithmic information content of the sequence $x, T(x), ..., T^n(x) \in X^*$ at accuracy ϵ and with respect to the interpretation I as:*

$$\mathcal{E}^I(x, n, \epsilon) = \min \left\{ |p| \ s.t. \ U(p) \in X^{n+1}, \max_{0 \leq i \leq n} (d(U_i(p), T^i(x))) < \epsilon \right\}.$$

As before we consider the behavior for $n \to \infty$ and define $\mathcal{E}^I(x, \epsilon) : X \times \mathbf{R} \to \frac{\mathbf{R}^*}{\simeq}$ as:

$$\mathcal{E}^I(x, \epsilon) = [J(\mathcal{E}^I(x, n, \epsilon))].$$

Remark 28 $\mathcal{E}^I(x, \epsilon)$ *is a non increasing function with respect ϵ.*

Finally, like in the definitions of initial data sensitivity we consider the behavior when ϵ goes to 0 and we define $\mathcal{E}^I(x) : X \to \mathcal{R}$ as

Definition 29 *The orbit complexity of x with respect to the interpretation I is defined as:*

$$\mathcal{E}^I(x) = \sup_{\epsilon \in \mathbf{R}^+} \mathcal{E}^I(x, \epsilon).$$

We now give some example of different behaviors of $\mathcal{E}^I(x)$. If x is a periodic point it is easy to see that $\mathcal{E}^I(x) \leq [J(log(n))]$. By the results of [10] and [17] it follows (see also Section 1) that if a system is compact, ergodic and has positive Kolmogorov entropy then for almost all points we have $\mathcal{E}^I(x) = [J(n)]$. We also remark that (when the space is compact) this is the maximum over all the possible behaviors. Indeed if X is compact, for each ϵ there is a finite cover made of balls with ideal center and radius ϵ, then a program that follows n steps of the orbit of any point with the accuracy ϵ can be simply made by listing n centers of the cover, then, if X is compact $\mathcal{E}^I(x) \leq [J(n)]$. In section 4 we will study the complexity of the orbits of another, less trivial example.

Lemma 30 *If I, J are computable interpretation functions from the same computable structure: $I, J \in \mathcal{I}$ then $\mathcal{E}^I(x) = \mathcal{E}^J(x)$. So the orbit complexity does not depend on the choice of the interpretation I in the computable structure $\mathcal{I}$ and we can define $\mathcal{E}^{\mathcal{I}}(x) = \mathcal{E}^I(x)$ for some $I \in \mathcal{I}$.*

Proof. Let us consider equivalent interpretations I_1, I_2 and $U = I_1(\mathcal{Q}(\mathcal{U}(*)))$ as in the definition of orbit complexity. Let us suppose that we have a minimal length program p_k for the interpretation

19

I_1 such that $\forall i < (k)$ we have $d(U_i(p_k), T^i(x)) < 2^{-\lambda-1}$, then there is a program p'_k for I_2 approximating the orbit of x with accuracy $2^{-\lambda}$ and $|p'_k| < |p_k| + c$. The program p'_k runs p_k finding strings s_i such that $I_1(s_i) = U_i(p_k)$, then using the equivalence between I_1 and I_2 it finds strings z_i such that $d(I_2(z_i), I_1(s_i)) < 2^{-\lambda-1}$ (Remark 11) by these strings it is easy to see how p'_k can approximate the orbit of x with accuracy $2^{-\lambda}$. It follows that $\mathcal{E}^{I_2}(x, 2^{-\lambda}) < \mathcal{E}^{I_1}(x, 2^{-\lambda+1})$, then we have $\mathcal{E}^{I_2}(x) \leq \mathcal{E}^{I_1}(x)$ and exchanging I_1 with I_2 we obtain the opposite inequality. $\square$

If X is compact then the orbit complexity does not depend not even on the computable structure.

Theorem 31 *If X is compact, if I is a computable interprepretation and J is another interpretation function (not necessarily computable) then $\mathcal{E}^I(x) \geq \mathcal{E}^J(x)$.*

Proof. Let $\epsilon > 0$, $s_1, ..., s_k \in \Sigma$ be a finite set of strings such that $B_\epsilon(I(s_1)), ..., B_\epsilon(I(s_k))$ is a cover of X. The set of strings is finite because X is compact. It is easy to see that there is an algorithm $A : \mathbf{N} \times \Sigma \to \{1, ..., k\}$ such that $A(i, s) = m$ implies that the $i - th$ point of $I(\mathcal{Q}(s)) \in B_\epsilon(I(s_m))$. That is: the algorithm gets a string and a natural number and outputs a set $B_\epsilon(I(s_m))$ of the cover in which the i-th point of the interpretation of the string as a sequence in X is contained. The algorithm calculates the distance between the $i - th$ point of $I(\mathcal{Q}(s)) \in B_\epsilon(I(s_m))$ and $I(s_z)$ for all $z \in \{1, ..., k\}$ with accuracy $\frac{\epsilon}{2}$, until it finds an s_m such that $d(U_i(s), I(s_m)) < \frac{\epsilon}{2}$ this is possible because I is a computable interpretation.

Now let us consider the interpretation J. Even for the interpretation J there is a finite set $\{s'_1, ... s'_{k'}\}$ such that $B_\epsilon(J(s'_1)), ..., B_\epsilon(J(s'_{k'}))$ is a cover of X. Now let us consider a function $G : \{1, ..., k\} \to \{1, ..., k'\}$ such that $G(i) = j$ if $I(s_i) \in B_\epsilon(J(s'_j))$. Being a function between finite sets G is a recursive function.

Now let p be a minimal length program that allows to follow the orbit of x for n steps with accuracy ϵ and interpretation I, that is

$$\max_{0 \leq i \leq n} (d(U_i(p), T^i(x))) < \epsilon, |p| = \mathcal{E}^I(x, n, \epsilon).$$

For each i by calculating $A(i, \mathcal{U}(p))$ we can find an m such that $T^i(x) \in B_{2\epsilon}(I(s_m))$ and then by function G we can find a j such that $T^i(x) \in B_{3\epsilon}(J(s_j))$.

Summarizing, this procedure allows (given the program p) to calculate a sequence of strings $s_{j_1}, ..., s_{j_n}$ such that we can follow the orbit of x with the interpretation J, for n steps and accuracy 3ϵ. This implies that $|p| + c \geq \mathcal{E}^J(x, n, 3\epsilon)$ where c is the length of the above procedure

20

and does not depend on n. From this we have $\mathcal{E}^I(x, \epsilon) \geq \mathcal{E}^J(x, 3\epsilon)$ and $\mathcal{E}^J(x) \leq \mathcal{E}^I(x)$. $\square$

From the above theorem we see the curious fact that in the compact case the orbit complexity reaches its maximum over all interpretations at a computable interptetation (a sort of Kolmogorov-Sinai theorem if we keep in mind the parallelism between orbit complexity and entropy) and the orbit complexity with respect to a computable structure does not depend on the choice of the computable structure (if some computable structure exists on the space). Moreover, all this is true independently of the properties of T.

Corollary 32 *If X is compact, if I and J are computable interpretations (not necessarily from the same computable structure) then $\mathcal{E}^I(x) = \mathcal{E}^J(x)$.*

The orbit complexity is invariant for constructive isomorphisms of dynamical systems over non compact spaces, it stated in the following propositions. As before we remark that if the space is compact constructivity is not required. We omit the proofs that are similar to the previous ones.

Theorem 33 *If (X, d, T), (Y, d', T') are topological dynamical systems over metric spaces with computable structures $\mathcal{I}, \mathcal{J}$ and f is onto and it is a morphism between $(X, d, \mathcal{I})$ and $(Y, d', \mathcal{J})$ such that the following diagram*

$$
\begin{array}{ccc}
 & f & \\
X & \rightarrow & Y \\
T \downarrow & & \downarrow \; T' \\
X & \rightarrow & Y \\
 & f &
\end{array}
\tag{2}
$$

commutes,if $x \in X$ and $y = f(x) \in Y$ then $\mathcal{E}^I(x) \geq \mathcal{E}^J(y)$.

Theorem 34 *If (X, d, T), (Y, d', T') are topological dynamical systems over compact metric spaces with computable structures $\mathcal{I}, \mathcal{J}$, if f is an homeomorphism $X \rightarrow Y$ such that the diagram 2 commutes and $x \in X$ and $y = f(x) \in Y$ then $\mathcal{E}^I(x) = \mathcal{E}^J(y)$.*

3.3 Complexity of points

Now we define a function $S^I(x, \epsilon) : X \times \mathbf{R} \rightarrow \mathbf{R}$, the function is a measure of the complexity of the points of X. The function is non increasing and measures how much information is necessary to approximate a given point of X with accuracy ϵ. Thus it is a function

21

that does not depend on the dynamics. In [16] a definition of local entropy for points of metric spaces was based on this idea and connections between S and the concept of dimension are shown. In particular $S(x, \epsilon)$ is related to the local dimension of X at x.

Definition 35 *If I is an interpretation function, $\mathcal{U}$ an universal computer we define the information contained in the point x with respect to the accuracy ϵ as:*

$$S^I(x, \epsilon) = \min\{|p| \quad s.t. \quad d(I(\mathcal{U}(p)), x) < \epsilon\}. \tag{3}$$

The function S depends on the interpretation I. In the following we will avoid to mention explicitly the superscript I when it is clear from the context. The function S depends also on the choice of $\mathcal{U}$. As stated in section 2.3 this function can be extended to a function $S^* : X \times \mathbf{R}^* \to \mathbf{R}^*$. Unfortunately S^* may be not compatible with the relation $\simeq$, for this reason we define $\overline{S'} : X \times \mathbf{R}^*/ \simeq \to \mathcal{R}$ as follows:

$$\overline{S}'(x, \alpha) = \sup_{a \in \alpha} [S^*(x, a)].$$

If $\alpha \in \mathcal{R}$ is an equivalence class then $\overline{S}'(\alpha)$ is the equivalence class of the supremum value of $S^*(a)$ where a ranges in the class α. Since an equivalence class in $R^*/ \simeq$ does not change by the adding of a constant, then the function S' does not depends more on the choice of the universal computer $\mathcal{U}$ in the definition of S. In the same way we define $\underline{S}' : X \times \mathbf{R}^*/ \simeq \to \mathcal{R}$ as:

$$\underline{S}'(x, \alpha) = \inf_{a \in \alpha} [S^*(x, a)].$$

Finally we extend $\overline{S}'$ and $\underline{S}'$ to functions $\overline{S}, \underline{S} : X \times \mathcal{R} \to \mathcal{R}$ as follows: if $a \in \mathcal{R}$ and $a = p(a_i)$ (the a_i are in $R^*/ \simeq$) then we define

$$\overline{S}(x, a) = \overline{p}(\overline{S}'(x, a_i)), \; \underline{S}(x, a) = \overline{p}(\underline{S}'(x, a_i))$$

this is well defined because $\mathcal{R}$ is closed by countable inf and sup and it does not depend on the choice of a_i in the class a. Because S is monotonic and then $(a_i) \approx (a_i')$ implies $S'(x, a_i) \approx S'(x, a_i')$.

If I and J are in the same computable structure $\mathcal{I}$ then the functions $\overline{S}^I$ and $\overline{S}^J$ are equal. $\overline{S}^I$ does not depend on the choice of the interpretation in the computable structure.

Lemma 36 $\overline{S}$ *and* $\underline{S}$ *are independent of the choice of* $I \in \mathcal{I}$*, in other words* $I, J \in \mathcal{I}$ *implies* $\overline{S}^I(x, \epsilon) = \overline{S}^J(x, \epsilon)$ *and* $\underline{S}^I(x, \epsilon) = \underline{S}^J(x, \epsilon)$.

22

Proof. The proof is very similar to the proof of Lemma 30 and we omit it. $\square$

We remark that if $X = \mathbf{R}^n$ or X is a finite dimensional manifold then $\overline{S} = \underline{S}$.

Remark 37 *If $(X, d, \mathcal{I})$ is a metric space with computable structure $\mathcal{I}$ and the lower box counting dimension [6] of X is finite : $\underline{\dim}_B(X) = d$ then $S(x, \delta) \leq -d \log \delta + C$ where C is a constant not depending on x and δ. Hence $\overline{S}$ and $\underline{S}$ coincides for all the points of X.*

The proof follows from the observation that if X is finite dimensional, then the minimum number n_ϵ of balls in a cover of X with radius ϵ is such that $n_\epsilon \sim \epsilon^{-d}$. By the computable structure we can construct the centers of a suitable cover with $n_\epsilon \sim \epsilon^{-d}$ and obtain that each point of X is approximated with accuracy ϵ by indicating a particular center of the cover, which costs $\leq log(n_\epsilon) + c \leq -d \log(\epsilon) + C$ bits, where c represents the length of the procedure that construct the centers of the suitable $\epsilon-$cover.

Lemma 38 *If $x \in \mathbf{R}^n$, μ is the Lesbegue measure on $\mathbf{R}^n$, if $\epsilon \to 0$ then for μ-almost all $x \in \mathbf{R}^n$, $S^I(x, \epsilon) = -nlog(\epsilon) + o(log(\epsilon))$ and if $a = \inf_\epsilon [J(a_{n,\epsilon})]$ then $\underline{S}(x, a) = \overline{S}(x, a) = \inf_\epsilon [J(log(a_{n,\epsilon}))]$.*

Proof. We prove that for almost all $x \in \mathbf{R}^n$ $\lim_{\epsilon \to 0} \frac{S^{\mathcal{I}}(x, \epsilon)}{-nlog(\epsilon)} = 1$. Theorem 12 of [16] states that the set

$$W^d = \{x \in \mathbf{R}^n \ s.t. \ \liminf_{i \to \infty} \frac{S^{\mathcal{I}}(x, 2^{-i})}{i} \leq d\} \qquad (4)$$

has Hausdorff dimension less or equal than d. This implies that if $\epsilon = 2^{-\gamma}$ then $\liminf_{\epsilon \to 0} \frac{S^I(x, \epsilon)}{-nlog(\epsilon)} = \liminf_{\gamma \to \infty} \frac{S^I(x, 2^{-\gamma})}{n\gamma}$. If $\liminf_{\epsilon \to 0} \frac{S^I(x, \epsilon)}{-nlog(\epsilon)} \leq 1$ then $x \in W^n$ because $S(x, 2^{-\gamma-1}) \leq S(x, 2^{int(-\gamma)}) \leq S(x, 2^{-\gamma+1})^7$. This implies that the set of the x s.t. $\liminf_{\epsilon \to 0} \frac{S^I(x, \epsilon)}{-nlog(\epsilon)} < 1$ is included in the set $\bigcup_{d < n} W^d$ that is a 0 measure set. To prove the other inequality it is enough to remark that each $x \in \mathbf{R}$ can be approximated with accuracy ϵ by specifying its first $-int(log(\epsilon))$ digits, if $x \in \mathbf{R}^n$ we need $-nlog(\epsilon)$ digits to explicit the n coordinates. $\square$

---

[6]See e.g. [13].
[7]$int(r)$ denotes the integer part of r

3.4 Initial data sensitivity and orbit complexity

Now we are ready to state the first proposition linking orbit complexity to initial data sensitivity.

Proposition 39 *If (X, T) is a dynamical system on a space with a computable structure $\mathcal{I}$, $I \in \mathcal{I}$ and T is constructive. There are constants c_1 and c_2 such that For all $x \in X, n \in \mathbf{N}, \epsilon \in \mathbf{R}^+$*

$$\mathcal{E}^I(x, n, 2\epsilon) < S^I(x, r(x, n, \epsilon)) + log(n) + c_1 \qquad (5)$$

$$S^I(x, R(x, n, 3\epsilon)) \leq \mathcal{E}^I(n, x, \epsilon) + c_2. \qquad (6)$$

Proof of 5. We will see that there is a program p such that $d(U_i(p), T^i(x)) < 2\epsilon, \forall i \leq n$ and $|p| \leq S(x, r(x, n, \epsilon)) + \log(n) + c$. If we have a program p_0 such that $d(I(\mathcal{U}(p_0)), x) \leq r(x, k, \epsilon)$ then $\forall i \leq k \ d(T^i(x), T^i(I(\mathcal{U}(p_0)))) < \epsilon$.

The idea is that by constructivity if we have the string $s_0 = \mathcal{U}(p_0)$ we can follow the orbit of $I(s_0)$ by an algorithm $A(s_0, k, \epsilon)$ (see Lemma 13). The program p will codify the following procedure:

1) run the program p_0 and compute $s_0 = \mathcal{U}(p_0)$
2) compute $s_i = A(s_0, i, \epsilon), \forall 1 \leq i \leq k$
3) compute the single string s such that $Q(s) = (s_0, ..., s_k)$.

The length of this program will be a constant (the above stated procedure) plus $\log(k)$ (the length of a binary representation of k) plus the length of p_0. If p_0 was supposed to be the shortest program such that $d(I(U(p_0)), x) < r(x, k, \epsilon)$ then its length is the value of $S(x, r(x, n, \epsilon))$ and the first part of the statement is proved.

Proof of 6. Let $e \in \mathbf{N}$ such that $2^{-e} < \epsilon$. If we have a program p such that $d(T^i(x), U_i(p_0)) < \epsilon$ for $0 < i < k$ we can find a string s s.t. $I(s) \in B(x, n, 3\epsilon)$ with the following procedure:

1) By p_0 compute the number k and the strings $s_0, ..., s_k$ such that $(s_{0,...,}s_k) = Q(\mathcal{U}(p_0))$
2) for each $c \in \Sigma$ do the following things: { compute $A(c, i, 2^{-e})$ for each $0 \leq i \leq k$, if for all $0 \leq i \leq k \ D(A(c, i, 2^{-e}), s_i, e + 2) < 2^{-e}$ then $s = c$ and stop the procedure. }

The procedure must stop because of the density of the image of I. At some time the step 2) will be computed with a c such that $I(c) \in B(x, n, 2^{-e-1})$ and this string will verify 2). On the other hand if we find a c that stops the procedure then it is easy to see that $I(c) \in B(x, n, 3\epsilon)$. This will implies that $d(x, I(c)) < R(x, k, 3\epsilon)$. Summarizing we have described a procedure that starting from a program

p_0 outputs a string s such that $d(I(s), x) < R(x, k, 3\epsilon)$. The code for this procedure will be a program containing p_0 and its length will be $p_0 + C$ where C represents the length of the code for the above procedure which does not depend on x and k, and the statement is proved.□

From the previous statement we obtain a relation between the indicators of orbit complexity and sensitivity.

Theorem 40 *If (X, T) is a dynamical system on a space with a computable structure $\mathcal{I}$ and T is constructive. For all $x \in X$*

$$\mathcal{E}^{\mathcal{I}}(x) \leq max(\overline{S}^{\mathcal{I}}(x, r(x)), [J(log(n))]) \tag{7}$$

$$\underline{S}^{\mathcal{I}}(x, R(x)) \leq \mathcal{E}^{\mathcal{I}}(x). \tag{8}$$

Proof. If we apply the homomorphism J to equation 5 we obtain

$$J(\mathcal{E}^I(x, n, 2\epsilon)) \leq J(S^I(x, r(x, n, \epsilon))) + J(log(n))$$

, then $J(\mathcal{E}^I(x, n, 2\epsilon)) \leq S^{*I}(x, J(r(x, n, \epsilon))) + J(log(n))$, and considering the equivalence classes:

$$\mathcal{E}^I(x, 2\epsilon) \leq max(\overline{S'}^I(x, r_\epsilon(x)), [J(log(n))]).$$

This is true for each ϵ, then $\mathcal{E}^I(x) \leq max(\overline{S}^I(x, r(x)) + [J(log(n))])$. As proved before all this equivalence classes does not depend on the choice of $I \in \mathcal{I}$ and we have Equation 7. In the same way we can obtain equation 8.□

By Lemma 38 for almost all points in $\mathbf{R}^n$ the function S is the logarithm, this, combined with proposition 39 implies the following formulas:

Theorem 41 *If $T : \mathbf{R}^n \to \mathbf{R}^n$ is constructive on $\mathbf{R}^n$ with the standard computable structure, for almost all $x \in X$*

$$\mathcal{E}^{\mathcal{I}}(x) \leq max(\inf_{\epsilon \in \mathbf{R}^+} ([J(log(r(x, n, \epsilon)))]), [J(log(n))]) \tag{9}$$

$$\inf_{\epsilon \in \mathbf{R}^+} [J(log(R(x, n, \epsilon)))] \leq \mathcal{E}(x) \tag{10}$$

As a corollary of Theorem 40 we can obtain the following interesting result: the set where the sensitivity to initial conditions is more than exponential in all directions has 0 Hausdorff dimension.

Theorem 42 *If (X,T) is a dynamical system on a compact metric space with a computable structure $\mathcal{I}$ and T is constructive. Then the set*

$$\overline{exp} = \{x \in X \ s.t. \forall h \in \mathbf{R}^+ \ R(x) < [J(2^{-hn})]\}$$

has zero Hausdorff dimension.

Proof. Let us consider a point x such that $R(x) < [J(2^{-hn})]$ by theorem 40 we know that $\underline{S}^{\mathcal{I}}(x, R(x)) \leq \mathcal{E}^{\mathcal{I}}(x)$ since X is compact we have $\mathcal{E}^{\mathcal{I}}(x) \leq [J(n)]$. Then $\underline{S}^{\mathcal{I}}(x, R(x)) \leq [J(n)]$. Since $\underline{S}$ is a non increasing function then $\underline{S}^{\mathcal{I}}(x, [J(2^{-hn})]) \leq \underline{S}^{\mathcal{I}}(x, R(x)) \leq [J(n)]$. Then by definition $\underline{S}^{\mathcal{I}}(x, [J(2^{-hn})]) = \inf_{\alpha \in [J(2^{-hn})]} [S^*(x, \alpha)] \geq [S^*(x, J(2^{-(h-\epsilon)n}))]$ for some small ϵ and then $[S^*(x, J(2^{-(h-\epsilon)n}))] \leq [J(n)]$, by this, setting $h' = h - \epsilon$ it follows that there is a bounded constant $c_1 \neq 0$ such that $S^*(x, J(2^{-h'n})) \leq c_1 J(n)$.

Let us consider the following set $A^d = \{x \in X | J(\frac{S(x, 2^{-i})}{i}) \leq d\}$, since for each sequence b_i we have $J(b_i) \geq \liminf(b_i)$ then $A^d \subset W^d$ where W^d is the set defined in eq. 4. Since the Hausdorff dimension of W^d is greater or equal than d then also $\dim_H(A^d) \leq d$. Now let us consider the set

$$A^{d,h'} = \{x \in X | J(\frac{S(x, 2^{-h'i})}{i}) \leq d\}$$

.

If $x \in A^{d,h'}$, let us set $k = h'i$ and let us consider $\frac{S(x, 2^{-k})}{\frac{k}{h'}}$, if $J(\frac{S(x, 2^{-h'i})}{i}) \leq d$ then $\liminf_{k \to \infty} (\frac{S(x, 2^{-k})}{\frac{k}{h'}}) \leq J(\frac{S(x, 2^{-h'i})}{i}) \leq d$ because $\frac{S(x, 2^{-h'i})}{i}$ is a subsequence of $\frac{S(x, 2^{-k})}{\frac{k}{h'}}$ (Axiom 16). Then $x \in W^{\frac{d}{h'}}$. This implies that if $x \in \overline{exp}$ then $x \in \cap_{h'} W^{\frac{d}{h'}}$ which has 0 Hausdorff dimension (again by [16] Theorem 12). $\square$

4 Applications to the Mannevile maps.

In this section, in order to give a non trivial example of application of the theory exposed in the previous sections we present some example of weakly chaotic dynamics. We construct a class of examples of dynamical systems over the unit interval with stretched exponential sensitivity to initial conditions and information content of the orbits that increases as a power law. We precise that the maps T_z we are going to study are not weakly chaotic in the sense of [29] (zero topological entropy), conversely they have positive topological entropy. In

this examples however for almost all the points (for the Lesbegue measure) the dynamics are weakly chaotic (low orbit complexity, low initial data sensitivity). Then we can say that the system is weakly chaotic with respect to the Lesbegue measure.

The examples are *piecewise linear* version of the Manneville map $T : [0, 1] \to [0, 1]$ defined as $T_z(x) = x + x^z \ (mod \ 1), \ z \in (1, \infty)$. The so called Manneville map comes from the theory of turbulence. It was introduced in [22] as an extremely simplified model of intermittent behavior in fluid dynamics, then the map was studied and applied in other areas of the physics (for example [1],[31],[27]).

The first study of the mathematical features of the Manneville map was done by Gaspard and Wang in [19]. However in our opinion In their paper some steps of the proofs were difficult to understand and some others were not rigorously formalized. In the following we outline the construction done in [19] for the study of the complexity of the piecewise linear Manneville maps by the theory of recurrent events [14]. Then we prove the main features of this important class of dynamical systems by the theory exposed in the previous sections. Another study of the Manneville map was done by C. Bonanno in [8] where the dynamics was studied also from a topological point of view. Notations: if $(a_i), (b_j) : \mathbf{N} \to \mathbf{R}$ are real sequences, in the following we will write $(a_i) \cong (b_j)$ if and only if $\frac{a_i}{b_i}$ is bounded, we also write $(a_i) \sim (b_j)$ if and only if $\lim_i \frac{a_i}{b_i} = 1$.

Let $\epsilon = (\epsilon_k) : N \to [0, 1]$ be a monotone sequence, such that $\lim_{k \to \infty} \epsilon_k = 0$. Let $T_\epsilon : [0, 1] \to [0, 1]$ be defined by:

$$
T_\epsilon(x) = \begin{cases} \frac{\epsilon_{k-2} - \epsilon_{k-1}}{\epsilon_{k-1} - \epsilon_k}(x - \epsilon_k) + \epsilon_{k-1} & k > 0, \epsilon_k < x \leq \epsilon_{k-1} \\ \frac{x - \epsilon_0}{1 - \epsilon_0} & \epsilon_0 < x \leq 1 \end{cases}
$$

To each sequence ϵ_i is then associated a piecewise linear map T_ϵ (see fig. 1) and a dynamical system $([0, 1], T_\epsilon)$.

In order to apply the theory of recurrent events, we associate to our dynamical system $([0, 1], T_\epsilon)$ a stochastic process X_i. The process is defined on the probability space (Ω, μ), where $\Omega = [0, 1]$ and μ is the Lesbegue measure as follows. Let us consider the sets $A_i, i \in \mathbf{N}$: $A_0 = (\epsilon_0, 1], ..., A_i = (\epsilon_i, \epsilon_{i-1}]$. Let $f : \mathbf{R} \to \mathbf{N}$ be given by $f(x) = n \Leftrightarrow x \in A_n$ (f associates to each point x the index of the set A_i in which x is included). The associated stochastic process is given by the random variables $X_i : \Omega \to \mathbf{R}$, given by $X_i(\omega) = f(T_\epsilon^i(\omega))$. As it was remarked in [19] the process X_i is a Markov chain with transition matrix
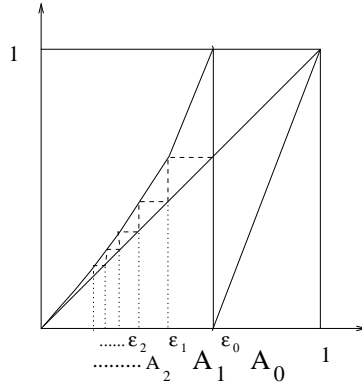
Figure 1: The map T_z and the partition A_i

$$\begin{pmatrix} p_0 & p_1 & p_2 & p_3 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$ where $p_i = \mu(A_i) = \mu(X_k = i | X_{k-1} =$

0). Summarizing: we constructed a family of dynamical systems (one for each infinitesimal sequence). To each one of them it is associated a Markov chain. The statistic behavior of the Markov chain can be studied by the theory of recurrent events [14] and will give information on the dynamics. The family $([0,1], T_\epsilon)$ is a large family of dynamical systems with different chaotic behavior. In the following we will study the systems in which $\epsilon_i \sim \frac{1}{i^\alpha}$. Such dynamical systems give rise to stretched exponential initial data sensitivity and power law orbit complexity. However if ϵ_i decreases slower than a power law ($\epsilon_i \sim \frac{-1}{log(i)}$ e.g.) we will have a variety of other possible behaviors of orbit complexity and initial data sensitivity that will be not studied here.

Now we consider a subset of the above family of systems, which we consider as P.L. versions of the Manneville map. Let $z \in (2, \infty)$, let us consider the map T_z associated by the above construction to the sequence

$$\epsilon_k = 1 - \sum_{i \geq k} \frac{1}{(i+1)^{\frac{z}{z-1}}}.$$

The map T_z is then associated to a Markov chain with transition

28

probabilities $p_0 = 1 - \epsilon_0$, $p_i = \frac{1}{(i+1)^{\frac{z}{z-1}}}$ $\forall i > 1$.

The theory of recurrent events can now give us information on the dynamics of the Markov chain. Let us consider the random variable $N_n(x)$ that is, the number of times that the event A_0 occurs until the time n: $N_n(x) = \#\{i|X_i(x) = 0\} = \#\{i|T^i(x) \in A_0\}$. Theorem 10 of [14] applied to our Markov chain says that if $1 - \sum_{i \geq k} p_i \sim A k^{-\alpha}$ then $E(N_n) \sim C n^\alpha$. Since for the Markov chain associated to T_z we have $1 - \sum_{i \geq k} p_i \sim \frac{1}{z-1} k^{\frac{1}{z-1}}$ then $E(N_n) \sim C n^{\frac{1}{z-1}}$.

Since the function T_z is not continuous at ϵ_0 our theory is not directly applicable to the dynamical system $([0,1], T_z)$. The following lemma and its proof shows how to extend our theory to discontinuous dynamical systems. A real number is constructive if it can be approximated at any accuracy by an algorithm.

Definition 43 *A number $z \in \mathbf{R}$ is said to be constructive if there is an algorithm $A_z(n) : \mathbf{N} \to \mathbf{Q}$ such that $A_z(n) = q$ implies $|q - z| < 2^{-n}$.*

The rational numbers, the algebraic numbers are constructive, and so are all the numbers that can be explicitly used for numerical purposes, for example π and e are constructive.

Lemma 44 *If z is constructive. For the dynamical system $([0,1], T_z)$, for all $x \in [0,1]$ equation 5 and 6 holds.*

proof. Let us consider the set $X = [0,1] - \{x|\exists k \in \mathbf{N}, T_z^k(x) = \epsilon_0\}$. Let us consider an interpretation function I' on X as follows: if I is defined as in equation 1 then $I' : \Sigma \to X$ is given by $I'(s) = I(s) + \pi$. I' is a computable interpretation because π is constructive and its image is in X because π is transcendent. If z is constructive then it is easy to see that T_z is constructive on $(X, d, \mathcal{I})$ where $I' \in \mathcal{I}$ then we can apply Proposition 39 to $(X, T_z, \mathcal{I})$. On the other hand, since the inclusion $i : X \to [0,1]$ is isometric the complexity of an orbit in X is equal to the complexity of the corresponding orbit in $[0,1]$: $\forall x \in X$ $\mathcal{E}^{I'}(([0,1], T_z), x, n, \epsilon) = \mathcal{E}^{I'}((X, T_z), x, n, \epsilon)$. If conversely $x \in [0,1] - X$ then $\mathcal{E}^{I'}(([0,1], T_z), x) = [J(log(n))]$ because the orbit of x converges to a fixed point. Then eq. 5 and 6 holds for all points of $[0,1]$ for the interpretation I'. Now, since $[0,1]$ is compact (see theorem 31) we have that the orbit complexity does not depend on I' and the statement is proved. $\square$

Now let us give an estimation of the initial data sensitivity of the map T_z by the behaviour of N_n. Let $x, y \in [0,1], d(x,y) = \Delta x(0) = \epsilon_0$. Since the derivative of T_z exists for almost all points and it is greater than 1, if ϵ_0 is small enough (e.g. is such that $\epsilon_0 \leq \frac{p_0}{4}$) then

29

$d(T_z(x), T_z(y)) \geq \epsilon_0$. If ϵ_0 is small enough (e.g. ϵ_0 is such that $\forall i \leq n, d(T_z^i(x), T_z^i(y)) \leq \frac{p_0}{4}$), then $d(T_z^n(x), T_z^n(y)) = \Delta x(n) \geq \epsilon_0 2^{cN_n}$ where c is a constant that depends on the derivative of T_z in A_0 and A_1.

If $d(x, y) = \epsilon_0 = \epsilon 2^{-cN_n}$ then $\Delta x(n) \geq \epsilon_0 2^{cN_n(x)} \geq \epsilon$. This implies that $y \notin B(n, x, \epsilon) = \{y \in X : d(T^i(y), T^i(x)) \leq \epsilon \ \forall \ 0 \leq i \leq n\}$ (as defined in section 3.1) and symmetrically $y' = 2x - y \notin B(n, x, \epsilon)$, then $\epsilon_0 = \epsilon 2^{-cN_n(x)} \geq R(n, \epsilon, x)$ (and then $R(x) \leq [J(2^{-cN_n(x)})]$).

By this we can give an estimation of the complexity of the orbits of the Manneville maps. In analogy with the results of [19] we are ready to state the following proposition about the asymptotic behavior of the average orbit complexity in the Manneville map.

Proposition 45 *If z is constructive. The average asymptotic behavior of the complexity of the orbits of T_z is for each ϵ*

$$E(\mathcal{E}^I(x, n, \epsilon)) \cong n^\alpha$$

where $\alpha = \frac{1}{z-1}$.

Proof. Since we have $R(n, \epsilon, x) \leq \epsilon 2^{-cN_n(x)}$ by proposition 39 and proposition 38 $-log(R(x, n, 3\epsilon)) + o(log(R(x, n, 3\epsilon))) \leq \mathcal{E}^I(x, n, \epsilon) + c_1$ for almost all $x \in [0, 1]$, then $\mathcal{E}^I(x, n, \epsilon) \geq cN_n(x) + o(N_n(x)) + const + log(\epsilon)$ and $E(\mathcal{E}^I(x, n, \epsilon)) \geq cE(N_n(x)) + E(o(N_n(x))) + const + log(\epsilon)$. But since $E(N_n(x)) \sim n^\alpha$ by the theory of recurrent events, we have that for all ϵ, $n^\alpha = O(E(\mathcal{E}^I(x, n, \epsilon)))$ (that implies $[J(E(\mathcal{E}^I(x, n, x, \epsilon)))] \geq [J(n^\alpha)]$ with our notations).

Now let us give an estimation from above of the average orbit complexity in the Manneville maps. As before we see that the number $N_n(x)$ will give the main part of the complexity of the orbit of x.

Let us consider an interpretation I for the standard computable structure in $[0, 1]$ and a minimal cover $B_\epsilon(I(s_0)), ..., B_\epsilon(I(s_l))$ of $[0, 1]$ (i.e. a cover of $[0, 1]$ such that each its proper subset does not cover $[0, 1]$). The number l of balls in this cover is then bounded by $l \leq \frac{1}{\epsilon}$. Now let us consider the sets A_i defined above. We remark that if $1 - \sum_{i \geq k} \frac{1}{(i+1)^{\frac{z}{z-1}}} < \epsilon$ then $A_k \subset B_\epsilon(I(s_0))$. Now we remark that the symbolic dynamics of the point x with respect to $\{A_i\}_{i \in \mathbf{N}}$ i.e. the sequence of $A_{i_0}, ..., A_{i_n}$ such that $T_z^n(x) \in A_{i_n}$ it is determined by the recurrence times of the set A_0. The sequence $(A_{i_n}) = A_{i_0}, ..., A_{i_k}$ must be such that if $i_n > 0$ then $A_{i_{n+1}} = A_{i_n-1}$ and if $i_n = 0$ then $A_{i_{n+1}}$ can be any one of the A_i (see fig 1). For example a possible sequence is $A_0, A_1, A_0, A_3, A_2, A_1, A_0, A_0, A_2, ...$ such a string is determined by the sequence P of numbers representing the recurrence times of A_0, i.e. for the above example $P = (0, 1, 3, 0, 2, ...)$. We remark that the

string $P_n(x) = (t_1, ..., t_{N_n(x)})$ representing the recurrence times of A_0 for n steps of the orbit of x contains $N_n(x)$ numbers. Then the binary length of $P_n(x)$ is about $\sum_{i=1}^{N_n} log(t_i)$.

An algorithm $A_\epsilon(n, m)$ to follow the orbit of x with accuracy ϵ (i.e. such that if $n \leq m$ and $A_\epsilon(n, m) = b$ then $T_z^n(x) \in B_\epsilon(I(s_b))$) can be constructed in the following way. The algorithm contains a string $P_m(x)$ of the recurrence times with respect to A_0, with this string we can reconstruct the symbolic orbit of x with respect to $\{A_i\}_{i \in \mathbf{N}}$ for m steps, as described above. Moreover the program contains another string $Q_m(x)$ containing at most $(l-1)N_n(x)$ numbers, each one is less or equal than l. The meaning of $Q_m(x)$ will be explained below.

The algorithm starts with a pointer to the first number of $Q_m(x)$. By P_m it calculates the set A_{i_n} such that $T_z^n(x) \in A_{i_n}$, if i_n is such that $1 - \sum_{j \geq i_n} \frac{1}{(j+1)^{\frac{z}{z-1}}} \leq \epsilon$ then $A_\epsilon(n, m) = 0$ because then $T_z^n(x) \in B_\epsilon(s_0)$. Else the algorithm outputs the number of $Q_m(x)$ indicated by the pointer and then set the pointer on the next number. In other words, when $T_z^n(x)$ may not be in $B_\epsilon(s_0)$ then the algorithm gets the number q such that $T_z^n(x) \in B_\epsilon(s_q)$ from the list $Q_m(x)$. This can not be too expensive because the point can be out of $B_\epsilon(s_0)$ at most $(l-1)N_n(x)$ times. The total length of the program implementing $A_\epsilon(n, m)$ for m steps of the orbit is then less or equal than $\sum_{i=1}^{N_n} log(t_i) + (l-1)N_n(x)log(l) + C$ where C is constant with respect to m, l depends on ϵ but not on m.

The term $\sum_{i=1}^{N_n} log(t_i)$ can be estimated as follows: we remark that while the random variables t_i are independent and identically distributed they have no finite expectation when $p_i \sim \frac{1}{i^\alpha}$, $1 < \alpha < 2$ ($E(t_i) = \sum ip_i$) instead the random variable $log(t_i)$ has finite expectation, let us say $E(log(t_i)) = \bar{t}$. Then $\sum_{i=1}^{N_n} log(t_i) = N_n(\frac{\sum_{i=1}^{N_n} log(t_i)}{N_n})$, by the law of large numbers we have that for each $\delta > 1$, for almost each x, eventually with respect to i we have $\frac{\sum_{i=1}^{N_n} log(t_i)}{N_n} < \delta \bar{t}$.

This implies that for almost each x, eventually with respect to n $\mathcal{E}^I(x, n, \epsilon) \leq cN_n(x)\delta\bar{t} + const$ and then $E(\mathcal{E}^I(x, n, x, \epsilon)) \geq CE(N_n(x)) + const$. But since $E(N_n(x)) \sim n^\alpha$ we have that for all ϵ, $E(\mathcal{E}^I(x, n, \epsilon)) = O(n^\alpha)$. $\square$

The above Theorem is an estimation of the average orbit complexity of the map T_z, to show how our theory can be applied to prove rigorously the statements of [19], however stronger results can be proved. By [19] page 4592 eq. 2.9 (which follows from [14] theorem 7, page 106) we have that there exists a constant A such that

$$\mu(\{N_n \geq \frac{n^\alpha}{Ax^\alpha}\}) \to G_\alpha(x)$$

where $\alpha = \frac{1}{z-1}$ and G_α is the Levi stable distribution law with param-

eter α. It follows that $N_n(x) \cong C n^\alpha$ for almost all points in the interval. From this, by the same proof as above it follows the pointwise estimation:

Theorem 46 *With the same notations as above, For almost all $x \in [0,1]$, $\mathcal{E}(x, n, \epsilon)) \cong n^\alpha$.*

Similar results are obtained by [8] using different techniques.

5   Numerical experiments

We want to remark that while the information content of an orbit (as it is defined in this work) is not computable (the algorithmic information content of a string is not a computable function) there is the possibility to have an empirical estimation of the quantity of information by the use of data compression algorithms. If instead to measure the information contained in a string by its algorithmic information content we consider as 'approximate' measure of the information content of the string the length of the string after it is compressed by a suitable coding procedure we obtain a computable notion of orbit complexity. In the positive entropy case the computable orbit complexity is a.e. equivalent to the previous one [18]. Such a definition of computable orbit complexity allows numerical investigations about the complexity of unknown systems. Unknown systems underlying for example some given time series or experimental datas.

The existence of a computable version of the orbit complexity motivates from the applicative point of view the study of the orbit complexity itself and its relations between the other measures of the chaotic behavior of a system.

In [3] and [23] such numerical investigations are performed by directly measuring the complexity of the orbits of the Manneville map and of the logistic map at the chaos threshold. The results agree with theoretical predictions and some interesting conjecture arises.

6   References

References

[1] Allegrini P., Barbi M., Grigolini P., West B.J., *Dynamical model for DNA sequences*, Phys. Rev. E, vol.52 nr.5, 5281-5297 (1995).

[2] Allegrini P., Grigolini P., West B.J., *A dynamical approach to DNA sequences*, Phys. Lett. A, **211**, 217-222 (1996).

[3] Argenti F. , Benci V. , Cerrai P. , Cordelli A. , Galatolo S. , Menconi G. *Information and dynamical systems: a concrete measurement on sporadic dynamics.To appear in Chaos, Solitons and Fractals*

[4] Batterman R.; White H. *Chaos and algorithmic complexity* Found. Phys. 26 (1996), no. 3, 307–336.

[5] Bak, Per; Tang, Chao; Wiesenfeld, Kurt. *Self-organized criticality.* Phys. Rev. A (3) 38 (1988), no. 1, 364–374.

[6] Benci V. *Alcune riflessioni su informazione, entropia e complessità.* Modelli matematici nelle scienze biologiche, P.Freguglia ed., QuattroVenti, Urbino, 1998.

[7] Benci, V. *An algebraic approach to nonstandard analysis.* Calculus of variations and partial differential equations (Pisa, 1996), 285–326, Springer, Berlin, 2000.

[8] Bonanno, C;The Manneville map: topological, metric and computational approach work in preparation.

[9] Brin M., Katok A. *On local entropy* in Geometric dynamics (vol 1007,Lecture notes in Mathematics) Springer,Berlin (1983).

[10] Brudno A.A. *Entropy and the complexity of the trajectories of a dynamical system* Trans. Moscow Math. Soc. **2** 127-151 (1983)

[11] Chaitin G.J. *Information, randomness and incompleteness. Papers on algorithmic information theory.* World Scientific, Singapore 1987.

[12] Cont, Rama *Modeling economic randomness: statistical mechanics of market phenomena.* Statistical physics on the eve of the 21st century, 47–64, Ser. Adv. Statist. Mech., 14, World Sci. Publishing, River Edge, NJ, 1999.

[13] Falconer, Kenneth. *Fractal geometry. Mathematical foundations and applications.* John Wiley & Sons, Ltd., Chichester, 1990.

[14] Feller, William *Fluctuation theory of recurrent events.* Trans. Amer. Math. Soc. 67, (1949). 98–119.

[15] Ford J. *Directions in classical chaos* in Directions in chaos, Vol. 1, 1–16, World Sci. Publishing, Singapore, 1987.

[16] Galatolo, S. *Pointwise information entropy for metric spaces.* Nonlinearity 12 (1999), no. 5, 1289–1298.

[17] Galatolo, S. *Orbit complexity by computable structures.* Nonlinearity 13 (2000), no. 5, 1531–1546.

[18] Galatolo, S. *Orbit complexity and data compression* Submitted to Discrete and Continuous Dynamical Systems.

[19] Gaspard P., Wang X.-J., *Sporadicity: between periodic and chaotic dynamical behaviors*, Proc. Natl. Acad. Sci. USA **85**, 4591-4595 (1988).

[20] Khinchin A.I. *Mathematical foundations of Information Theory* Dover Publications , New York.

[21] Yang, Y; Grigolini, P;em On the time evolution of the entropic index, Phys. Lett. A 263, 323-330 (1999).

[22] Manneville P., *Intermittency, self-similarity and 1/f spectrum in dissipative dynamical systems*, J. Physique **41**, 1235-1243 (1980).

[23] Menconi,G; Bonanno,C *Computational information for the logistic map at the chaos threshold* work in preparation

[24] Meson, A. ; Vericat,F. *Invariant of dynamical systems: a generalized entropy* J. Math. Phys. 37, 4480 (1996).

[25] Petersen K. *Ergodic Theory* Cambrigdge University Press,Cambridge (1983)

[26] Pesin Y.B. *Characteristic Lyapunov exponents and smooth ergodic theory* Russ. Math. Surv. 32 (1977)

[27] Pollicott M., Weiss H. *Multifractal analysis of Lyapunov exponent for continued fraction and Manneville-Pomeau transformations and applications to Diophantine approximation.* Comm. Math. Phys. 207 (1999), no. 1, 145–171.

[28] M.B. Pour-El and J.I. Richards *Computability in Analysis and Phisics* Springer, Berlin (1988).

[29] Smital, J. Chaotic functions with zero topological entropy. Trans. Amer. Math. Soc. 297 (1986), no. 1, 269–282.

[30] Takens, Floris; Verbitski, Evgeny. Generalized entropies: Renyi and correlation integral approach. Nonlinearity 11 (1998), no. 4, 771–782.

[31] Toth, T.I.; Liebovitch, L.S. *Models of ion channel kinetics with chaotic subthreshold behaviour.* Z. Angew. Math. Mech. 76, Suppl. 5, 523-524 (1996).

[32] Tsallis, C.; Plastino, A. R.; Zheng, W.-M.*Power-law sensitivity to initial conditions—new entropic representation.* Chaos Solitons Fractals 8 (1997), no. 6, 885–891.

[33] White H. *Algorithmic complexity of points in dynamical systems* Ergodic Theory Dynam. Syst. **13** 807-830 (1993)

[34] Zvorkin A.K., Levin L.A. *The complexity of finite objects and the algorithmic-theoretic foundations of the notion of information and randomness* Russ. Math. Surv. 25 (1970)